

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.

THIS PAGE BLANK (USPTO)

DEVICE, METHOD AND SYSTEM FOR PATTERN RECOGNITION

Patent number: JP2000075887

Publication date: 2000-03-14

Inventor: ASANO KOJI; MINAMINO KATSUKI; HONDA HITOSHI;
HAMADA OSAMU

Applicant: SONY CORP

Classification:

- international: G10L15/06; G06K9/62; G06K9/68

- european:

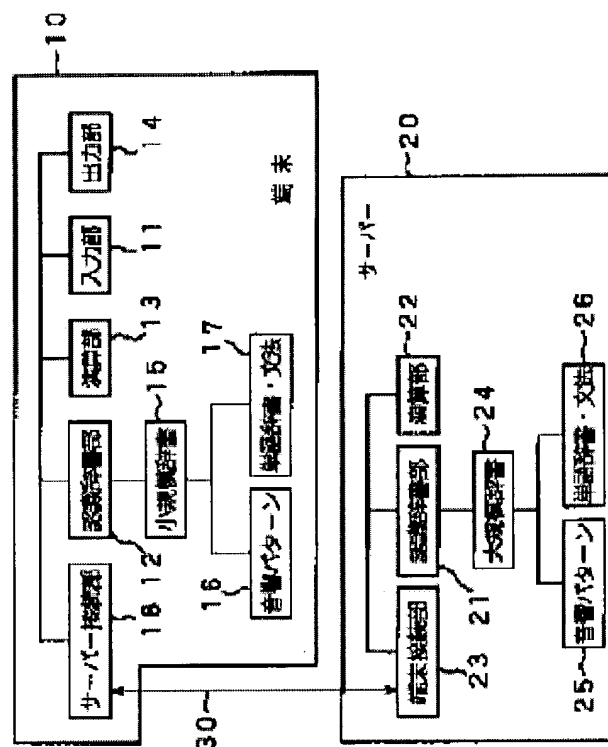
Application number: JP19980246401 19980831

Priority number(s):

Abstract of JP2000075887

PROBLEM TO BE SOLVED: To increase the recognition object patterns and to reduce the network load.

SOLUTION: In the pattern recognition system, a terminal 10, which is provided with a small scale dictionary 15 that becomes a reference for pattern recognition, and a server 20, which is provided with a large scale dictionary 24 having a larger scale than the dictionary 15, are connected by a network. The terminal 10 conducts a recognition process for the pattern inputted into an input section 11 using the dictionary 15. If the recognition result is not appropriate, recognition is conducted by the dictionary 24 of the server 20.



Data supplied from the esp@cenet database - Worldwide

THIS PAGE BLANK (USPTO)

Family list

1 family member for:

JP2000075887

Derived from 1 application.

1 DEVICE, METHOD AND SYSTEM FOR PATTERN RECOGNITION

Publication info: **JP2000075887 A** - 2000-03-14

Data supplied from the **esp@cenet** database - Worldwide

THIS PAGE BLANK (USPTO)



PATENT ABSTRACTS OF JAPAN

JPA 2000-075887

(11) Publication number: **2000075887 A**(43) Date of publication of application: **14.03.00**

(51) Int. Cl.

G10L 15/06**G06K 9/62****G06K 9/68**(21) Application number: **10246401**(22) Date of filing: **31.08.98**(71) Applicant: **SONY CORP**

(72) Inventor: **ASANO KOJI**
MINAMINO KATSUKI
HONDA HITOSHI
HAMADA OSAMU

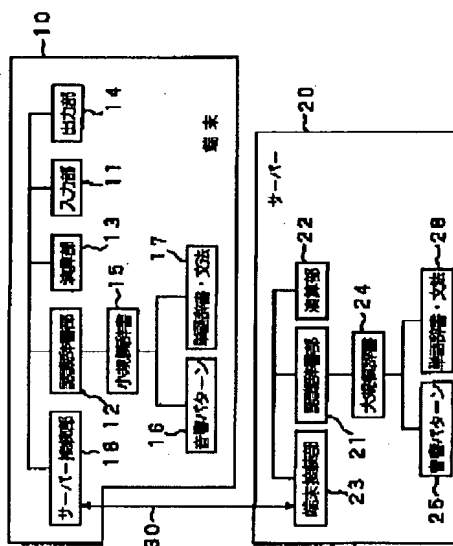
**(54) DEVICE, METHOD AND SYSTEM FOR PATTERN
 RECOGNITION**

(57) Abstract:

PROBLEM TO BE SOLVED: To increase the recognition object patterns and to reduce the network load.

SOLUTION: In the pattern recognition system, a terminal 10, which is provided with a small scale dictionary 15 that becomes a reference for pattern recognition, and a server 20, which is provided with a large scale dictionary 24 having a larger scale than the dictionary 15, are connected by a network. The terminal 10 conducts a recognition process for the pattern inputted into an input section 11 using the dictionary 15. If the recognition result is not appropriate, recognition is conducted by the dictionary 24 of the server 20.

COPYRIGHT: (C)2000,JPO



THIS PAGE BLANK (USPTO)

(51) Int.Cl. ⁷	識別記号	F I	テマコード (参考)		
G10L 15/06		G10L 3/00	521	V	5B064
G06K 9/62	650	G06K 9/62	650	C	5D015
9/68		9/68		B	9A001

審査請求 未請求 請求項の数16 O L (全11頁)

(21) 出願番号 特願平10-246401

(22) 出願日 平成10年8月31日 (1998.8.31)

(71) 出願人 000002185

ソニー株式会社

東京都品川区北品川6丁目7番35号

(72) 発明者 浅野 康治

東京都品川区北品川6丁目7番35号 ソニー株式会社内

(72) 発明者 南野 活樹

東京都品川区北品川6丁目7番35号 ソニー株式会社内

(74) 代理人 100067736

弁理士 小池 晃 (外2名)

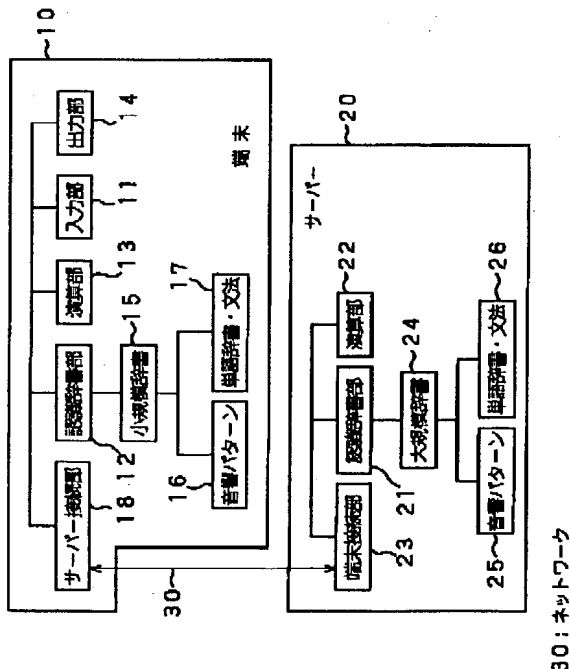
最終頁に続く

(54) 【発明の名称】 パターン認識装置、方法及びシステム

(57) 【要約】

【課題】 認識対象パターンの増大とネットワーク負荷の低減を実現する。

【解決手段】 パターンの認識に用いる基準となる小規模辞書15を備える端末10と、パターン認識に用いる基準となり、小規模辞書15より規模の大規模な大規模辞書24を備えるサーバ20とがネットワークで接続されてなるパターン認識システムであって、端末10は、入力部11に入力したパターンについて、小規模辞書15で認識を行い、その認識の結果が不適切であるとサーバ20の備える大規模辞書20で認識を行う。



【特許請求の範囲】

【請求項1】 パターンを入力する入力手段と、
 パターンの認識に用いる基準となる第1の辞書と、
 パターンの認識に用いる基準となる辞書であって上記第1の辞書より規模の大きい第2の辞書と、
 パターンを認識する認識処理手段と、
 パターンの認識の結果が適切であるか否かを判断する判断処理手段と、

パターンの認識結果を出力する出力手段と、
 上記入力手段に入力されたパターンを上記認識処理手段において上記第1の辞書を用いて認識し、上記認識処理手段における認識結果を上記判断処理手段にて判断し、
 上記判断処理手段で上記認識結果が適切であると判断されると上記出力手段で上記認識結果を出力し、上記判断処理手段で上記認識結果が適切であると判断されないと、
 上記パターンを上記認識処理手段で上記第2の辞書を用いて判断し、上記認識結果を上記出力手段で出力するように制御する制御手段とを有することを特徴とするパターン認識装置。

【請求項2】 上記判断処理手段は、上記認識処理手段における認識結果が適切か否かをパターン認識の過程で得られる尺度を用いて自動的に判断することを特徴とする請求項1記載のパターン認識装置。

【請求項3】 上記認識処理手段は、上記尺度として信頼性尺度を採用することを特徴とする請求項2記載のパターン認識装置。

【請求項4】 上記第1の辞書には、上記第2の辞書にて認識に利用されたパターンが登録されることを特徴とする請求項1記載のパターン認識装置。

【請求項5】 上記第2の辞書にて認識が行われたパターンの上記第1の辞書への登録は、上記第2の辞書への参照の頻度に応じて自動的に行われることを特徴とする請求項4記載のパターン認識装置。

【請求項6】 上記第1の辞書は、初期状態ではパターンが登録されていないことを特徴とする請求項4記載のパターン認識装置。

【請求項7】 上記認識処理手段は、音声、画像又は文字の少なくとも一つを認識の対象とすることを特徴とする請求項1記載のパターン認識装置。

【請求項8】 パターンを入力する入力手段と、
 パターンの認識に用いる基準となる第1の辞書と、
 パターンの認識に用いる基準となる辞書であって上記第1の辞書より規模の大きい第2の辞書と、
 パターンを認識する認識処理手段と、
 データを出力する出力手段と、
 パターンの認識結果の確認を受け付ける確認手段と、
 上記入力手段に入力されたパターンを上記認識処理手段において上記第1の辞書を用いて認識し、上記認識処理手段における認識結果を上記判断処理手段にて判断し、
 上記判断処理手段における判断結果を上記確認手段に提

示し、上記確認手段にて上記認識結果について確認を受けると、上記出力手段で上記認識結果を出力し、上記確認手段にて上記認識結果について確認が受けられないと上記認識処理手段において上記第2の辞書を用いて認識し、上記出力手段で上記認識結果を出力するように制御する制御手段とを有することを特徴とするパターン認識装置。

【請求項9】 上記制御手段は、上記提示手段に、上記認識処理手段における認識処理結果とともに、上記認識処理手段におけるパターン認識の過程で得られる尺度を提示することを特徴とする請求項8記載のパターン認識装置。

【請求項10】 上記認識処理手段は、上記尺度として信頼性尺度を採用することを特徴とする請求項9記載のパターン認識装置。

【請求項11】 上記第1の辞書には、上記第2の辞書にて認識に利用されたパターンが登録されることを特徴とする請求項8記載のパターン認識装置。

【請求項12】 上記第2の辞書にて認識が行われたパターンの上記第1の辞書への登録は、上記第2の辞書への参照の頻度に応じて自動的に行われることを特徴とする請求項11記載のパターン認識装置。

【請求項13】 上記第1の辞書は、初期状態ではパターンが登録されていないことを特徴とする請求項11記載のパターン認識装置。

【請求項14】 パターン認識に用いる基準となる第1の辞書と、パターンの認識の基準となる辞書であって、上記第1の辞書より規模の大きい第2の辞書とを用い、パターンを入力する入力工程と、

パターンを認識する認識処理工程と、
 パターンの認識の結果が適切であるか否かを判断する判断処理工程と、

パターンの認識結果を出力する出力工程と、
 上記入力工程に入力されたパターンを上記認識処理工程において上記第1の辞書を用いて認識し、上記認識処理工程における認識結果を上記判断処理工程にて判断し、
 上記判断処理工程で上記認識結果が適切であると判断されると上記出力工程で上記認識結果を出力し、上記判断処理工程で上記認識結果が適切であると判断されないと、
 上記パターンを上記認識処理工程で上記第2の辞書を用いて判断し、上記認識結果を上記出力工程で出力するように制御する制御工程とを有することを特徴とするパターン認識方法。

【請求項15】 パターン認識に用いる基準となる第1の辞書と、パターン認識に用いる基準となる辞書であって、上記第1の辞書より規模の大きい第2の辞書とを用い、

パターンを入力する入力工程と、
 パターンを認識する認識処理工程と、
 データを出力する出力工程と、

パターンの認識結果の確認を受け付ける確認工程と、上記入力工程に入力されたパターンを上記認識処理工程において上記第1の辞書を用いて認識し、上記認識処理工程における認識結果を上記判断処理工程にて判断し、上記判断処理工程における判断結果を上記確認工程に提示し、上記確認工程にて上記認識結果について確認を受けると、上記出力工程で上記認識結果を出力し、上記確認工程にて上記認識結果について確認が受けられないと上記認識処理工程において上記第2の辞書を用いて認識し、上記出力工程で上記認識結果を出力するように制御する制御工程とを有することを特徴とするパターン認識方法。

【請求項16】 パターンを入力する入力手段と、パターンの認識に用いる基準となる第1の辞書と、上記第1の辞書を用いてパターンを認識する第1の認識処理手段と、パターンの認識の結果が適切であるか否かを判断する判断処理手段とパターンの認識結果を出力する出力手段と、パターンの認識の制御を行う制御手段とを有する端末と、パターンの認識に用いる基準となる辞書であって、上記第1の辞書より規模の大きい第2の辞書と、上記第2の辞書を用いてパターン認識処理を行う第2の認識処理手段とを有するサーバとがネットワークによって結ばれてなるパターン認識システムであって、上記制御手段は、上記入力手段に入力されたパターンを上記第1の認識処理手段において上記第1の辞書を用いて認識し、上記第1の認識処理手段における認識結果を上記判断処理手段にて判断し、上記判断処理手段で上記認識結果が適切であると判断されると上記出力手段で上記認識結果を出力し、上記判断処理手段で上記認識結果が適切であると判断されないと、上記パターンを上記第2の認識処理手段で上記第2の辞書を用いて判断し、上記認識結果を上記出力手段で出力するように制御することを特徴とするパターン認識システム。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】 本発明は、パターン認識を利用した装置及び方法に関し、特にネットワークで接続された複数の機器におけるパターン認識装置、方法及びシステムに関する。

【0002】

【従来の技術】 一般に、パターン認識を行う場合には、照合する際に基準になるパターンを登録したパターン辞書が大きくなるほど、より詳細な照合を行うことができ、認識精度を高めたり、またより多くの種類のパターンの中からの認識を行うことが可能となる。

【0003】 しかし一方で、パターン辞書を大きくすると、認識処理に必要となる演算量や記憶容量が大きくな

る。

【0004】 そのため、通常の装置においても、認識処理のレスポンスが悪化することが避けられない。

【0005】 また、特に携帯を想定した小型端末など演算量や記憶容量の制約が厳しい場合には、大規模なパターン辞書を利用したパターン認識を端末単体で実現するのは困難であった。

【0006】 このような小型の端末で大規模なパターン辞書を利用したパターン認識処理を行う方法としては、特開平7-152797号公報で開示された明細書中に示されているように、ネットワークを介して、サーバに接続し、サーバにおいて大規模なパターン辞書を利用したパターン認識処理を行い、その結果をネットワークを介して端末に送ると行った方法が提案されている。

【0007】 すなわち、上記明細書中には、図4に示すように、情報アクセスシステムの一構成例が開示されている。この情報アクセスシステムは、センターシステム101およびそれと、例えばPSTNや統合サービスディジタル網(integrated services digital network; ISDN)などの電話回線などで接続され、センターシステム101と通信可能なターミナル102とから構成されている。

【0008】 ターミナル102は、センターシステム101の音声入力型データベース検索・送出システム103をインターフェイスとして、例えば図書館情報、新聞情報、雑誌情報、新刊情報などの刊行物の情報が記憶された図書データベース105にアクセスし、所望する情報を得ることができるようになされている。なお、図書データベース105には、刊行物の情報として、そのテキストデータなどの他、その音韻情報および韻律情報からなる発音記号情報が記憶されており、ターミナル102では、これらの情報を受信し、刊行物のテキストデータの表示、印刷の他、発音記号情報から合成音を生成し、これにより刊行物を朗読した音声を得ることができるようになされている。

【0009】 センターシステム101は、音声入力型データベース検索・送出システム103、オーサリングシステム104、図書データベース105、およびネットワーク制御部106から構成されている。オーサリングシステム104は、刊行物の情報を、図書データベース5に記憶させる、刊行物の情報のデータベース化や、図書データベース105に記憶された情報を編集することができるようになされている。ネットワーク制御部106は、センターシステム101と他のセンターシステムとの通信を制御し、これによりセンターシステム101の図書データベース5と、他のセンターシステムにおけるデータベースとの共有化を図ることができるようになされている。

【0010】

【発明が解決しようとする課題】 しかし、大規模なパタ

ーン辞書を用いたパターン認識処理には、処理のレスポンスという問題がある。

【0011】すなわち、認識用のパターン辞書を大規模化すると、入力パターンの照合に時間がかかり、認識結果がでるまでの時間が長くなり、レスポンスが悪くなる。そのため機器の操作性が悪くなることが考えられる。

【0012】例えば、パターン認識処理の一例である音声認識処理を用いた端末を考えてみる。端末をコントロールするためのコマンドなど良く使われる語の認識のレスポンスが悪いと、端末の操作性が大きく損なわれる。

【0013】また、図4に示した情報アクセスシステムのような大規模なパターン辞書を用いたパターン認識処理を、端末とサーバを用いた実現するような構成のパターン認識処理装置の場合、次のような問題がある。

【0014】すなわち、端末においてパターン認識を実行しようとする、端末はまずネットワークを介してサーバに接続し、入力パターンのデータを転送し、サーバでパターン認識処理を行い、その結果をふたたびネットワーク経由でサーバから得るという処理を行う必要がある。そのため、パターン認識結果を得るのに時間が掛かり、より一層のレスポンスの低下が起こる。

【0015】さらに、サーバにおいて常にパターン認識を行う場合、パターン認識を行う度にネットワークを利用した端末とサーバを接続しなければならず、ネットワークでの負荷が増え、特に携帯端末などで常にネットワークに接続されているわけではない環境では、ネットワーク利用のコストが増大する。

【0016】本発明は、上述の実情に鑑みてなされるものであって、パターン認識処理のレスポンス短縮と認識対象パターンの増大を実現するようなパターン認識装置及び方法を提供することを目的とする。

【0017】

【課題を解決するための手段】上述の課題を解決するために、本発明に係るパターン認識装置は、パターンを入力する入力手段と、パターンの認識に用いる基準となる第1の辞書と、パターンの認識に用いる基準となる辞書であって上記第1の辞書より規模の大きい第2の辞書と、上記入力手段に入力されたパターンについて上記第1の辞書で認識を行い、その結果が不適切であると上記第2の辞書にて認識を行う認識手段と、上記認識手段からのパターンの認識の結果を出力する出力手段とを有するものである。

【0018】また、本発明に係るパターン認識方法は、パターンを入力する入力工程と、パターンの認識に用いる基準となる第1の辞書、及びパターンの認識に用いる基準となる辞書であって上記第1の辞書より規模の大きい第2の辞書を用い、上記入力工程に入力されたパターンについて上記第1の辞書で認識を行い、その結果が不適切であると上記第2の辞書にて認識を行う認識工程

と、上記認識工程からのパターン認識の結果を出力する出力工程とを有するものである。

【0019】そして、本発明に係るパターン認識システムは、パターンの認識に用いる基準となる第1の辞書を備える端末と、パターンの認識に用いる基準となる辞書であって上記第1の辞書より規模の大きい第2の辞書を備えるサーバとがネットワークにて結ばれてなるパターン認識システムであって、上記端末は、上記端末に入力したパターンについて、上記第1の辞書で認識を行い、その結果が不適切であると上記サーバの備える第2の辞書で認識を行うものである。

【0020】すなわち、本発明ではパターン認識における照合処理の基準となるパターン辞書を、小規模な第1の辞書と大規模な第2の辞書の2つに分けて保持し、それを必要に応じて使い分ける。

【0021】すなわち、まず小規模な第1の辞書を用いて、パターン認識処理を行う。その結果が、適切な認識結果が得られれば、パターン認識処理を終了する。もし適切な結果が得られなかった場合は、大規模な第2の辞書を用いてふたたびパターン認識処理を行う。

【0022】

【発明の実施の形態】以下、本発明の実施の形態について、図面を参照して詳細に説明する。

【0023】まず、本発明の第1の実施の形態として、パターン認識により音声を認識する音声認識装置について説明する。

【0024】音声認識装置は、図1に示すように、ユーザによりデータの入出力の操作がなされる端末10と、大容量のデータベースを備えてデータの処理を集中して行うサーバ20とから構成されている。これら端末10およびサーバ20は、情報を伝送する路線であるネットワーク30により接続が可能である。

【0025】まず、端末10について説明する。端末10は、ユーザによるデータの入出力を受け付け、ネットワーク30を介してサーバ20と接続するものである。

【0026】端末10は、音声の入力を受け付ける入力部11と、パターン辞書に相当する認識辞書部12と、音声パターンを認識辞書部12により照合する演算部13と、演算部13から送られた結果を出力する出力部14とを有している。

【0027】入力部11は認識対象となる音声パターンの入力を受け付ける部分で、マイク、アンプ、AD変換器などで構成される。入力部11では、入力された音声パターンをデジタル化して演算部13に送る。

【0028】認識辞書部12は、パターン辞書に相当するものである。認識辞書部12は、半導体メモリやハードディスクなどで構成される。認識辞書部12は、音声認識の場合は、一般に認識の基本単位となる音響パターンと、音響パターンがどのように連鎖し得るかを規定する単語辞書、文法から構成される。

【0029】本実施の形態においては、認識辞書部12は、小規模な認識辞書である小規模辞書15に実装されている。小規模辞書15は、音響パターン16および単語辞書・文法17から構成されている。

【0030】なお、本実施の形態には、小規模辞書より規模の大きい認識辞書である大規模辞書も備えられている。この大規模辞書は、小規模辞書15より精細な音響パターン、大語彙の単語辞書、複雑な文法を備えている。後述するように、この大規模辞書はサーバ20に備えられるので、端末10からはネットワーク30により接続することができる。

【0031】演算部13はCPU、半導体メモリなどで構成される。演算部13は、入力部11より送られた音声パターンと、認識辞書部12に格納されている音響パターンを、単語辞書、文法に規定されたパターンの接続条件を満たす形で照合を行う。演算部13は、このような照合による音声認識により得られた音声認識結果を出力部14に送る。

【0032】出力部14は、ディスプレイや音声合成装置などから構成される。出力部は、演算部13より送られた音声認識結果をユーザに提示する。

【0033】また、端末10は、ネットワーク30に接続するインターフェース部であるサーバ接続部18を有している。サーバ接続部18はモデムなどで構成され、必要に応じてネットワークを経由してサーバと接続する。

【0034】続いて、サーバ20について説明する。サーバ20は、大容量のデータベースを備え、ネットワーク30を介して端末から接続されることにより、データ処理を集中して行うものである。

【0035】サーバ20は、パターン辞書に相当する認識辞書部21と、音声パターンを認識辞書部21を用いて認識する演算部22と、ネットワーク30に接続するインターフェースである端末接続部23とを有している。

【0036】認識辞書部21は、端末10に備えられる認識辞書部12と略々同じ構成である。すなわち、認識辞書部21は、パターン辞書に相当するものであって、半導体メモリやディスクなどで構成される。ただし、端末10に備えられる小規模辞書15より規模の大きい大規模辞書24が備えられている。大規模辞書24には、小規模辞書15より精細な音響パターン25、大語彙の単語辞書、複雑な文法26が備えられている。

【0037】演算部22は、端末10に備えられる演算部13と略々同じ構成である。すなわち、演算部22は、CPU、半導体メモリなどで構成され、端末10からネットワークを経由して送られた音声パターンと、認識辞書部21に格納されている音響パターンを、単語辞書、文法に規定されたパターンの接続条件を満たす形で照合を行う。ただし、サーバ20に備えられる演算部2

2には、端末10に備えられる演算部13より高性能なものが用いられている。演算部22は、認識辞書部21を用いた照合による音声認識により得られた音声認識結果を端末10に送る。

【0038】端末接続部23はモデムなどで構成されている。端末接続部23は、端末10からのサーバ接続命令を受け、このサーバ接続命令に基づいて、端末10との間の接続を制御する。

【0039】頻繁に利用されるパターンに関しては、小規模な認識辞書である小規模辞書15に登録しておくことによって、頻繁に利用される語に関しては、小規模辞書15でレスポンスが早い認識処理を行うことができる。

【0040】さらに、大規模な認識辞書である大規模24辞書をサーバ20上に備え、ネットワーク30を介して利用することによって、より多くのパターンとより精細に照合した認識を行うことも可能である。しかも、これは必要となった場合にだけ実行されるので、常時ネットワーク30で接続してサーバ20で認識処理を行う場合に比べて、ネットワーク30の負荷や通信コストの低減が可能となる。

【0041】続いて、音声認識装置における音声認識処理の流れについて、図2に示すフローチャートを参照して説明する。

【0042】最初のステップS11においては、端末10の入力部11に対して、音声が入力される。入力部11では、入力音声をデジタル化し、演算部13に送る。

【0043】ステップS12においては、演算部13は、入力部から送られたデジタル化された音声を、端末10の認識辞書部12の小規模辞書15を用いて認識処理を行う。すなわち、小規模辞書12に備えられる音響パターン16および単語辞書・文法17を用いて認識処理を行う。

【0044】ステップS13においては、ステップS12にて行われた音声の認識処理の結果を判断する。認識処理の結果の判断を行うには、音声認識装置が自動で行う方法と、ユーザによる確認の方法との2つの方法が考えられる。

【0045】音声認識の結果を音声認識装置が自動で行う方法としては、一般に信頼性尺度(confidence measure;C)と呼ばれる認識結果の信頼性に関する尺度を基準として判断する方法がある。

【0046】ここで、信頼性尺度Cには、認識スコアが第1位の候補の認識スコアP(Wfirst)、認識スコアが第2位の候補の認識スコアP(Wsecond)について、次のように計算される値が用いられる。

$$C = P(Wfirst) - P(Wsecond)$$

また、信頼性尺度Cは、さらに第3位の候補の認識スコアP(Wthird)以下を用いて、

$$C = W(Pfirst) / (P(Wsecond) - P(Wthird))$$

d))

により与えることもできる。

【0047】上述の信頼性尺度は、一般に、パターン認識処理の結果、認識スコアが第1位の候補と、第2位以下の候補の認識スコアの差が大きいほど、認識結果の信頼性が高いと想定されることに基づいて定義したものである。

【0048】このような信頼性尺度を基準として判断する方法によると、音声認識装置は、音声の認識処理の過程で信頼性尺度を計算し、信頼性尺度の値が所定の閾値を超えた場合には認識結果がOKであると判断し、信頼性尺度の値が上記閾値より小さい場合には認識結果がNGであると判断する。

【0049】一方、ユーザによる確認による方法では、認識処理の結果を演算部13から出力部14に送る。そして、出力部14において、画面表示または音声による提示によってユーザに伝達し、その結果をユーザが判断して、認識結果の良し悪しを判断する。

【0050】ステップS14においては、ステップS13における認識結果の判断により分岐する。すなわち、認識結果がOKと判断された場合には、ステップS20において最終結果出力として、その認識結果を最終結果として、出力部14に送り、ユーザに出力して、処理は終了する。認識結果がNGであると判断された場合には、次のステップS15に進む。

【0051】ステップS15においては、端末10はサーバ接続部18からネットワーク20を介してサーバに接続する。そして、端末10は、音声認識に必要なデータをサーバ20に転送する。ここで、音声認識に必要なデータとしては、入力音声データそのものでもいいが、音響分析をした結果や、さらにベクトル量子化した結果などを転送するもできる。音声分析した結果やベクトル量子化した結果を送ることにより、転送する情報量を減らすことができる。サーバ20では、転送されたデータを端末接続部23で受け取り、そのデータを演算部22に送る。

【0052】ステップ16においては、サーバ20に演算部は、端末10から送られたデータを認識辞書部21の大規模辞書24に保持されているデータと照合して、認識を行う。そして、演算部22は、大規模辞書24を用いた音声認識の結果を端末接続部23からネットワーク30を介して端末10に送る。最後のステップS20においては、端末10は、サーバより送られてきた認識結果を最終結果として出力する。

【0053】なお、サーバ20による大規模辞書24を用いた認識時においては、以下の処理も同時に行うこともできる。

【0054】すなわち、S17において、大規模辞書24中の各パターン、特に単語認識に関しては、大規模辞書24を参照して認識結果を得る際に、過去に何回認識

結果として利用されているかを保持しておくようにする。そして、大規模辞書を参照した認識が行われる度に、認識結果として得られたパターンの利用回数に1を加える。

【0055】そして、ステップS18において、過去の利用回数がある閾値を超えた場合、またはある期間における利用回数の頻度がある閾値を越えた場合には、ステップS19において該当するパターンをサーバより端末に転送し、端末中の小規模辞書に登録する。

【0056】このように、頻繁にアクセスされるパターンを端末10の小規模辞書15に転送することによって、端末10がサーバ20にアクセスする回数を減らすことができる。さらに、認識処理のレスポンスの改善やネットワーク30の負荷の低減に効果がある。

【0057】なお、端末10の小規模辞書15へのパターンの登録は、上記のように音声認識装置により自動的に行うことも可能であるし、登録前にユーザの確認を取るよう構成することも可能である。

【0058】小規模辞書15には登録されていないが、大規模辞書24に登録されている単語で、ユーザがよく利用する単語に関しては、小規模辞書15にも登録することによって、大規模辞書24へのアクセス回数を減らすことができ、さらなるレスポンスとネットワーク30負荷の改善が可能となる。

【0059】続いて、音声認識装置の実際の処理の例について説明する。音声認識装置の実際の処理の例としては、端末10によるデータベース検索に上記のような音声認識装置を組み込むことが考えられる。例えば図書情報を検索する端末10と、図書情報を待つサーバ20を考える。図面のクリアやスクロールなど端末10を制御するコマンドなどを小規模辞書15に保持し、図書の名称や著者名、出版社名など図書情報に関連し、またその数も多いものを大規模辞書24に登録する。

【0060】ユーザは画面のクリアなど良く用いられるコマンドに関しては、端末10に内蔵の小規模辞書15による認識で、ネットワーク30を利用することなくすばやいレスポンスで端末を操作することが可能である。そして図書の名前など小規模辞書15に登録されていないものを入力した場合は、端末10は自動的にネットワーク30を経由してサーバ20に接続し、認識処理を実行するので、ユーザは認識用辞書が分散配置されていることやネットワークと接続されていることなどを意識することなく端末を利用することができる。

【0061】またこの時、端末10とサーバ20の間の通信のプロトコルを決めることによって、個々の機器の制御コマンドの名称などは各機器毎の独自に決めることが可能となる。

【0062】次に、本発明の第2の実施の形態として、パターン認識により手書き文字の認識を行う手書き文字認識装置について説明する。

【0063】まず、手書き文字認識装置の構成について、図3を参照して説明する。なお、この図3においては、図1に示した音声認識装置に対応する部分については同一の符号を附することにする。

【0064】音声認識装置は、ユーザによりデータの入出力の操作がなわれる端末10と、大容量のデータベースを備えてデータの処理を集中して行うサーバ20とから構成されている。これら端末10およびサーバ20は、情報を伝送する路線であるネットワーク30により接続が可能である。

【0065】端末10は、ユーザによるデータの入出力を受け付け、ネットワーク30を介してサーバ20と接続するものである。端末10は、音声の入力を受け付ける入力部11と、パターン辞書に相当する認識辞書部12と、音声パターンを認識辞書部12により照合する演算部13と、演算部13から送られた結果を出力する出力部14と、ネットワーク30を介してサーバ20と接続するサーバ接続部18とを有している。

【0066】認識対象となる手書き文字パターンの入力を受け付ける入力部11は、タッチパッドなどで構成される。入力部11では、入力された手書き文字パターンをデジタル化して演算部13に送る。

【0067】パターン辞書に相当する小規模辞書15を備える認識辞書部12は、認識の基本単位となる文字パターン19と、文字パターン19がどのように構成されるか規定する単語辞書・文法17を有している。

【0068】なお、端末10の他の部分の構成、およびこの端末10における認識の処理については、上述した音声認識装置と同様であるので説明を省略する。

【0069】サーバ20は、パターン辞書に相当する認識辞書部21と、音声パターンを認識辞書部21を用いて認識する演算部22と、ネットワーク30に接続するインターフェースである端末接続部23とを有している。

【0070】認識辞書部21に備えられる、小規模辞書14より規模の大きい大規模辞書24は、小規模辞書15より精細な文字パターン19、大語彙の単語辞書、複雑な文法26を備えている。

【0071】なお、サーバ20の他の部分の構成、およびこのサーバ20における認識の処理については、上述した音声認識装置と同様であるので説明を省略する。

【0072】このように、本実施の形態の手書き文字認識装置においては、小規模辞書15および大規模辞書24に音響パターンにかかわって文字パターンが認識されていること、および入力部11がタッチパッド構成され、入力された手書き文字をデジタル化することが、上述した第1の実施の形態の音声認識装置と相違している。

【0073】この手書き文字認識装置における処理手順は、音声パターンに代わって文字パターンを処理することをのぞくと、図2で示した音声認識装置における処理

手順と原則として同一である。

【0074】すなわち、手書き文字認識装置の処理工程を、図2によって説明すると、最初のステップS11では、端末10において、音声入力に代わって、入力部11のタッチパッドなどに手書き文字を入力する。次のステップS12においては、演算部13は、入力部11入力された手書き文字を認識辞書部12の小規模辞書15で文字パターン19および単語辞書・文法17を用いて認識する。

10 【0075】ステップS13においては、ステップS12における認識の結果を判断する。この認識結果の判断は、上述したように、例えば信頼性尺度 (confidence measure) を基準として自動で行う方法と、ユーザに提示して判断する方法とがある。認識結果の判断がOKとされると、ステップS20に進んで最終結果を出力部14に表示して終了する。認識結果の判断がNGと判断されると、ステップS15に進む。

【0076】ステップS15においては、端末10は、サーバ接続部18によりネットワーク30を介してサーバ30に接続し、入力された手書き文字をサーバ20に送る。ここで、サーバ20に送られる手書き文字のデータは、入力された手書き文字そのものである必要はなく、分析をした結果や、適当な量子化をした結果を送ることにより、転送に要する情報量を提言することができるのは、上述の音声認識装置と同様である。

【0077】ステップS16においては、サーバ20において、演算部22は、認識処理部21の備える大規模辞書24の文字パターン27および単語辞書・文法26を用いて、手書き文字を認識する。そして、サーバ20は、認識処理の結果を端末接続部23によりネットワーク30を介して端末10に送る。そして、ステップS20においては、端末は、サーバにて認識された結果を出力部14に表示して、一連の工程を終了する。

【0078】なお、認識結果の利用回数を計測し、アクセスに頻度により大規模辞書24に登録された単語を小規模辞書に登録することにより、端末10のサーバ20へのアクセス回数を提言することができるのも、上述した音声認識装置と同様である。このような大規模辞書24から小規模辞書15への単語の登録については、ステップS17からステップS19に示されているが、上述の第1の実施の形態と同様であるので、説明を省略することにする。

【0079】以上のように、手書き文字認識装置における処理手順は、音声認識装置における処理手順と、パターン認識用のアルゴリズムの処理が、音響パターン認識により行う音声認識用であるか、手書き文字認識による手書き文字認識用であるかが相違している。

【0080】続いて、この手書き文字認識装置における、実施の処理について説明する。

【0081】手書き文字認識装置においては、音声認識

装置ではユーザから音声によって情報を入手していたのだが、手書き文字による入力に変更される。これによって人名や地名などにある難しい漢字など、通常端末 10 の小規模辞書 15 には登録されていないような文字も、ネットワーク 30 を介して大規模辞書 24 に接続することによって認識することが可能となる。

【0082】なお、図 3 に示した手書き文字認識装置の構成においては、小規模辞書 15 と大規模辞書 24 を端末 10 とサーバ 20 に分散して配置した構成を示した。しかし、小規模辞書 15 と大規模辞書 24 の両方を同一の装置内に内蔵することも可能である。その場合、まず小規模辞書 15 による認識を行うことによって、頻度が

高く入力されるものはすばやいレスポンスで認識処理を行うことができる。

【0083】また、第 1 の実施形態の音声認識装置においても、小規模辞書 15 と大規模辞書 24 を同一の装置に備えることにより、頻度が高く入力されるものはすばやいレスポンスで認識処理を行うことができる。

【0084】以上説明したように、本実施の形態は、パターン認識の際に用いる照合用の辞書を小規模なもの

10

20

で、小型端末における認識対象パターンの増大とネットワーク負荷の低減を実現するものである。

【0085】

【発明の効果】本発明は、小規模なパターン辞書と大規模なパターン辞書を併用することによって、頻繁に利用されるパターンを小規模辞書に登録して、パターン認識のレスポンスを早めると同時に、大規模辞書へのアクセスの回数を減らすことが可能となる。これによって、大規模辞書を利用したパターン認識がネットワークを介して接続可能なサーバによって提供されている場合は、常時サーバを利用してパターン認識を行う場合に比べて、ネットワーク負荷の低減が可能になる。

【図面の簡単な説明】

【図 1】音声認識装置の概略的な構成を示すブロック図である。

【図 2】音声認識装置における処理手順を示すフローチャートである。

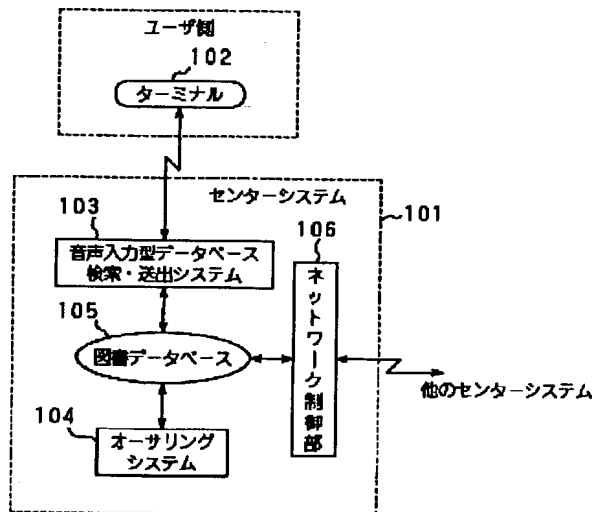
【図 3】手書き文字認識装置の概略的な構成を示すブロック図である。

【図 4】情報アクセスシステムの一例の構成を示すブロック図である。

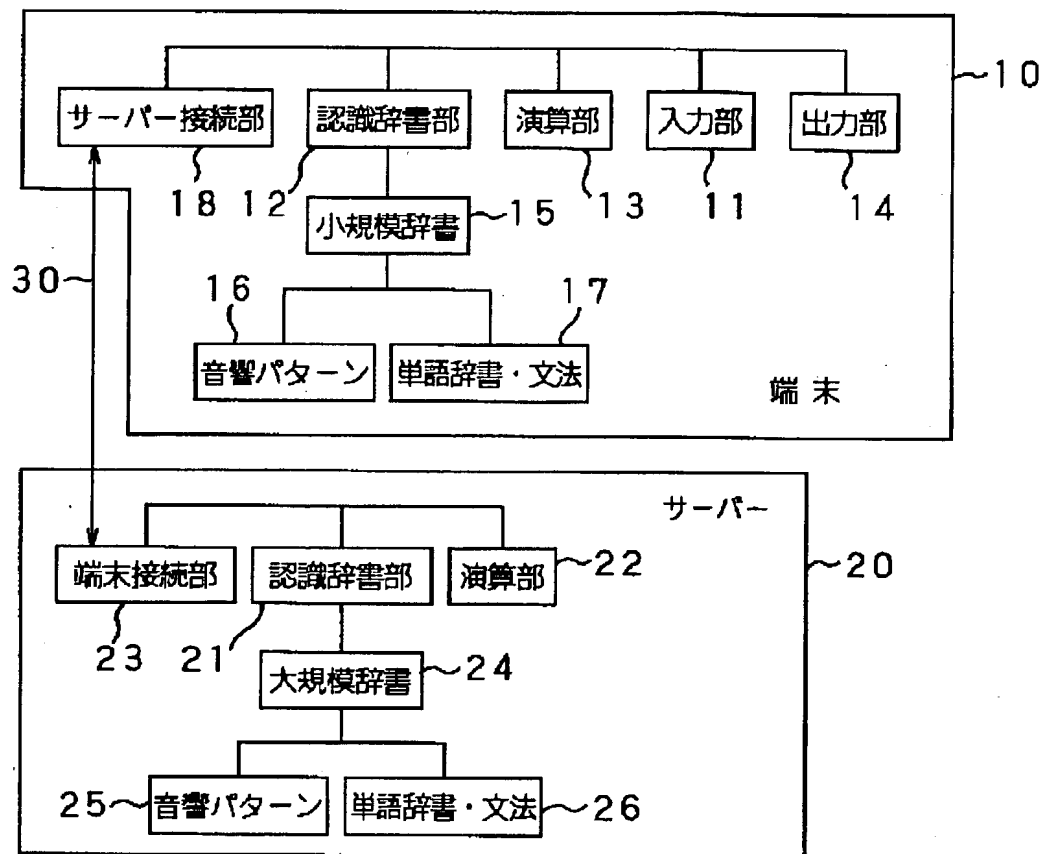
【符号の説明】

10 端末、15 小規模辞書、20 サーバ、24 大規模辞書、30 ネットワーク

【図 4】

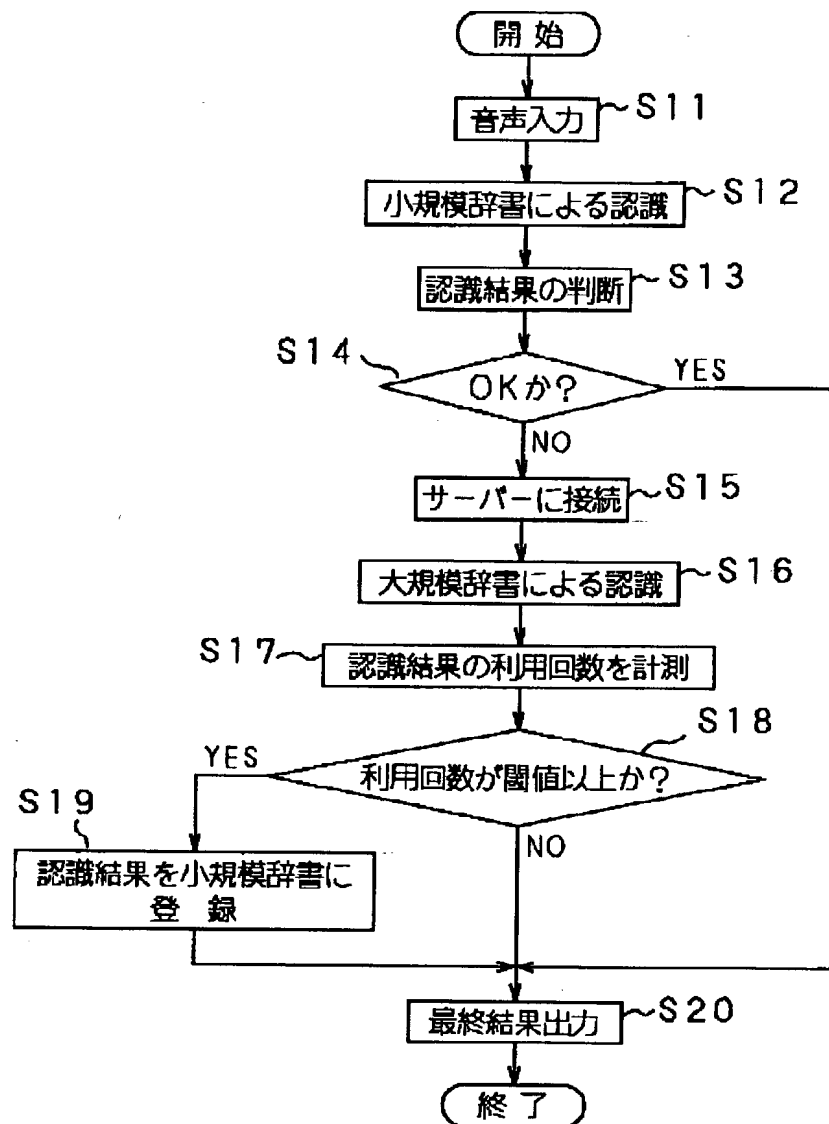


【図1】

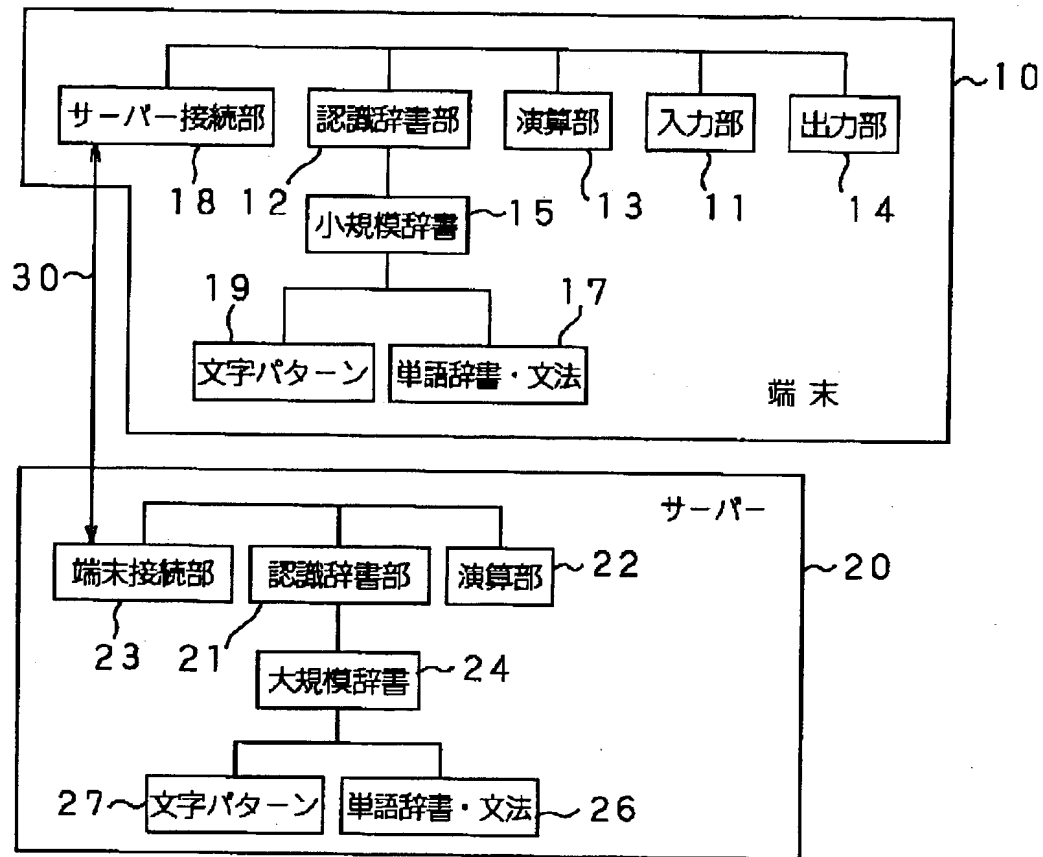


30: ネットワーク

【図2】



【図3】



30: ネットワーク

フロントページの続き

(72)発明者 本田 等
 東京都品川区北品川6丁目7番35号 ソニ
 ー株式会社内

(72)発明者 浜田 修
 東京都品川区北品川6丁目7番35号 ソニ
 ー株式会社内

Fターム(参考) 5B064 DA15 DA27 DA29
 5D015 HH13 LL10
 9A001 HH17 HH21 HH22

THIS PAGE BLANK (USPTO)